



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
[www.uspto.gov](http://www.uspto.gov)

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/915,603	07/26/2001	Alvin Charles Richardson	13018:19	8249
33697	7590	02/02/2005	EXAMINER	
GREGORY SCOTT SMITH			BASHORE, WILLIAM L	
P.O. BOX 88148			ART UNIT	PAPER NUMBER
ATLANTA, GA 30356			2176	
DATE MAILED: 02/02/2005				

Please find below and/or attached an Office communication concerning this application or proceeding.

<b>Office Action Summary</b>	<b>Application No.</b>	<b>Applicant(s)</b>	
	09/915,603	RICHARDSON ET AL.	
	<b>Examiner</b>	<b>Art Unit</b>	
	William L. Bashore	2176	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

### **Period for Reply**

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
  - If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
  - If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
  - Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

## Status

- 1)  Responsive to communication(s) filed on 10 November 2004.  
2a)  This action is **FINAL**.                  2b)  This action is non-final.  
3)  Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

## **Disposition of Claims**

- 4)  Claim(s) 1-30 is/are pending in the application.  
4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.

5)  Claim(s) \_\_\_\_\_ is/are allowed.

6)  Claim(s) 1-30 is/are rejected.

7)  Claim(s) \_\_\_\_\_ is/are objected to.

8)  Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

## **Application Papers**

- 9)  The specification is objected to by the Examiner.

10)  The drawing(s) filed on \_\_\_\_\_ is/are: a)  accepted or b)  objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)  The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

**Priority under 35 U.S.C. § 119**

- 12)  Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).  
a)  All b)  Some \* c)  None of:  
1.  Certified copies of the priority documents have been received.  
2.  Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.  
3.  Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

\* See the attached detailed Office action for a list of the certified copies not received.

**Attachment(s)**

- 1)  Notice of References Cited (PTO-892)  
2)  Notice of Draftsperson's Patent Drawing Review (PTO-948)  
3)  Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)  
Paper No(s)/Mail Date \_\_\_\_\_.  
  
4)  Interview Summary (PTO-413)  
Paper No(s)/Mail Date \_\_\_\_\_.  
5)  Notice of Informal Patent Application (PTO-152)  
6)  Other: \_\_\_\_\_.

**DETAILED ACTION**

1. This action is responsive to communications: amendment filed 11/10/2004, to the original application filed 7/26/2001, said application is a CIP of application 09/294,701 filed 4/19/1999 (now U.S. Patent No. 6,782,505).
2. The rejection of claim 20 under 35 USC 112 second paragraph, has been withdrawn as necessitated by amendment.
3. The rejection of claims 1-15 under 35 U.S.C. 101 as being directed to non-statutory subject matter, has been withdrawn as necessitated by amendment.
4. Claims 1-30 remain rejected under 35 U.S.C. 103(a) as being unpatentable over Huck and Weigel.
5. Claims 1-30 pending. Claims 1, 16 are independent.

*Specification*

6. The disclosure is objected to because of the following informalities page 1 of Applicant's specification should be amended to reflect that U.S. Patent Application Serial No. 09/294,701 is now U.S. Patent No. 6,782,505. Appropriate correction is required.

*Claim Rejections - 35 USC § 103*

7. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

8. **Claims 1-30 are rejected under 35 U.S.C. 103(a) as being unpatentable over Huck, G., et al. (hereinafter Huck), Jedi: extracting and synthesizing information from the Web, IEEE Cooperative Information Systems 1998, August 20-22, 1998, pp.32-41, in view of Weigel, A. et al. (hereinafter Weigel), Lexical postprocessing by heuristic search and automatic determination of the edit costs, IEEE Document Analysis and Recognition, 1995, August 14-16, 1995, pp.857-860.**

**In regard to independent claim 1,** Huck teaches a method of wrapper generation (JEDI) for extracting data from documents, said method comprising creating grammars for identifying patterns of symbols, said pattern containing prefix, value, and suffix patterns (Huck Abstract, also section 3 “Extraction Language”, second column, especially “<string><blanks><number><blanks><number>”). It is noted that Huck relies upon a combination of pattern matching (as explained above) with grammars in its implementation.

Huck teaches resolving ambiguities by exploring all possible solution, and then picking the best solution, therefore teaching identification and eventual selection of candidate matches, based upon ranking (Huck section 4 “Parsing Strategy”, fifth paragraph from top of said section, see also section 7 “Conclusion and Further Research”, especially second paragraph from top of said section).

Huck does not specifically teach determining a “cost” associated with choosing a “best” ranked solution (candidate match). However, Weigel teaches automatic determination of edit costs (Weigel page 857, Title, section 1 “Introduction” – especially at top of second column, also page 859 section 6 “Learning the values of y”). It would have been obvious to one of ordinary skill in the art at the time of the invention to apply the edit costs of Weigel to Huck’s ranking, providing Huck the benefit of increased accuracy in its ranked results.

**In regard to dependent claims 2, 3,** Huck does not specifically teach edit distances. However, Weigel teaches Insertions, substitutions, and deletions, which are used in calculating edit distances (operations) (see

Weigel page 858 sections 3 and 4, also page 859 – at top of first column; compare with claim 2). It would have been obvious to one of ordinary skill in the art at the time of the invention to apply Weigel to Huck, providing Huck the benefit of edit distance calculations for more accurate candidate selection.

Huck teaches an example pattern string (Huck section 3 “Extraction Language” – second column; compare with claim 3).

**In regard to dependent claims 4, 5,** Huck teaches generation of all possible solutions (spans of interest), as well as subjecting patterns/grammars as filters (Huck section 3 “Extraction Language” second paragraph from bottom of said section). Huck also teaches an example implementation of its invention comprising the query (typically involving keywords) of extracted data (Huck section 5 “Example”).

**In regard to dependent claims 6, 7, 8,** Huck does not specifically teach thresholds, or lowest cost selection. However, Weigel teaches thresholds, as well as calculations for determining lowest costs (Weigel page 859 section 5 “Speeding up the search”). It would have been obvious to one of ordinary skill in the art at the time of the invention to apply Weigel to Huck, providing Huck the benefit of edit distance calculations for more accurate candidate selection.

**In regard to dependent claims 9, 10, 11,** Huck does not specifically teach adjustments, or weights, or addition. However, Weigel teaches calculations for determining costs, as well as adjustments and weights (Weigel page 859 section 5 “Speeding up the search”, section 6 “Learning the values of y”, also page 858 section 3). It would have been obvious to one of ordinary skill in the art at the time of the invention to apply Weigel to Huck, providing Huck the benefit of various edit distance calculations for more accurate candidate selection.

**In regard to dependent claim 12,** claim 12 incorporates substantially similar subject matter as claimed in claim 1, and is rejected along the same rationale.

**In regard to dependent claims 13, 14,** Huck teaches generation of all possible solutions (spans of interest), as well as subjecting patterns/grammars (i.e. regular expressions) as filters (Huck section 3 “Extraction Language” second paragraph from bottom of said section). Huck also teaches an example implementation of its invention comprising the query (typically involving keywords) of extracted data (Huck section 5 “Example”).

**In regard to dependent claim 15,** claim 15 incorporates substantially similar subject matter as claimed in claim 1, and is rejected along the same rationale.

**In regard to claims 16-30,** claims 16-30 reflect the computer readable medium comprising computer executable instructions used for performing the methods as claimed in claims 1-15, respectively, and are rejected along the same rationale.

*Response to Arguments*

9. Applicant's arguments filed 11/10/2004 have been fully and carefully considered but they are not persuasive.

Applicant argues on page 9 of the amendment that the primary reference does not teach a candidate match for the three items listed in claim 1 (prefix pattern, suffix pattern, and value pattern). It is noted that Huck teaches extracting data from documents comprising creation of grammars for identifying patterns of symbols, said pattern containing prefix, value, and suffix patterns. The pattern sequence: “<string><blanks><number><blanks><number>” can be interpreted as a form of prefix/value/suffix pattern, and is processed accordingly using both pattern matching and grammars. Huck additionally teaches resolving

Art Unit: 2176

ambiguities by exploring all possible solution, and then picking the best solution, therefore teaching identification and eventual selection of candidate matches, based upon ranking.

Applicant argues on page 10 of the amendment that secondary reference (Weigel) does not teach determining a cost associated with each of the candidate matches, as well as lacking motivation, etc. It is noted that Weigel teaches automatic determination of edit costs. Since Weigel applies edit costs to compare edit strings, it would have been obvious to one of ordinary skill in the art at the time of the invention to apply the edit costs of Weigel to Huck's candidate ranking, providing Huck the benefit of increased accuracy in its ranked results. Both references deal with string pattern matching, and both try to find the best matching solution, therefore said reference have been combined accordingly. The publishing of the Weigel reference before the Huck reference in the same trade journal by a German company, by itself, does not preclude the combination of said references under 35 U.S.C. 103(a).

Applicant argues on pages 10-11 of the amendment that Huck does not teach "spans of interest", as well as keywords, as claimed by Applicant. It is noted that the examiner interprets Huck's teaching of all generated possible solutions as spans of interest (as well as subjecting patterns/grammars as filters) (Huck section 3 "Extraction Language" second paragraph from bottom of said section). Without further clarification within the relevant claim, all possible solutions can be interpreted as a span of interest, since Huck tries to achieve the best solution, all possible solutions would be of interest to Huck, accordingly. Huck also teaches an example implementation of its invention comprising the query (typically involving keywords) of extracted data (Huck section 5 "Example"). A keyword query acts as a form of filter (i.e. filtering out unwanted data until a match is found).

### *Conclusion*

10. **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

Art Unit: 2176

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

11. Any inquiry concerning this communication or earlier communications from the examiner should be directed to William L. Bashore whose telephone number is (571) 272-4088. The examiner can normally be reached on 11:30am - 8:00pm EST.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Joseph Feild can be reached on (571) 272-4090. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).



WILLIAM L. BASHORE  
PATENT EXAMINER  
TECH CENTER 2100  
January 28, 2005